RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10

Source:

Date Processed by STIC:

ENTERED



IFWP

RAW SEQUENCE LISTING DATE: 02/21/2006
PATENT APPLICATION: US/10/567,630 TIME: 12:58:40

Input Set : F:\39467A.txt

Output Set: N:\CRF4\02212006\J567630.raw

```
3 <110> APPLICANT: Alitalo, et al.
     5 <120> TITLE OF INVENTION: MATERIALS AND METHODS FOR COLORECTAL CANCER SCREENING,
DIAGNOSIS AND
             THERAPY
     8 <130> FILE REFERENCE: 28113/39467A
C--> 10 <140> CURRENT APPLICATION NUMBER: US/10/567,630
C--> 10 <141> CURRENT FILING DATE: 2006-02-08
    10 <150> PRIOR APPLICATION NUMBER: US 60/494,221
    11 <151> PRIOR FILING DATE: 2004-08-08
    13 <160> NUMBER OF SEQ ID NOS: 50
    15 <170> SOFTWARE: PatentIn version 3.2
    17 <210> SEO ID NO: 1
    18 <211> LENGTH: 49275
    19 <212> TYPE: DNA
    20 <213> ORGANISM: Homo sapiens
    22 <220> FEATURE:
    23 <221> NAME/KEY: misc_feature
    24 <223> OTHER INFORMATION: Prox-1 Genomic
    26 <400> SEQUENCE: 1
    27 ggaaatgaaa aaagaaaaag aaaaaaaaaa aaaaagacct gcgtcctgga agagctagtg
                                                                             60
    29 tgagccgggc gccgctcgcg ccgtctcccg ctttgcatag tgcccgcaga tggctcgctc
                                                                            120
    31 cggcccaggc gcggcgatcc agtcgtcccg aagctgtggc agggtggggg tggggtggg
                                                                            180
    33 ggccggggat ggaggccggg gagggggagc gcagggcccc tctcccctcc tctctccca
                                                                            240
    35 gcccctcacc cccacccctt ttatattttt ttttcctccc aagttctctt gccttgctat
                                                                            300
    37 ccccccttga atccgaaggc gcctcgcgat tgggtgctgg ggccgggtac gtcagataga
                                                                            360
    39 ctgtgacgtg cagtetteet gttteettea getgtgtett aaagtaaate ttgttgtgga
                                                                           420
    41 geggageeet cagetgaggg agegetetga aataatacae cattgeagee qqqqaaaqea
                                                                           480
    43 gageggegea aaagagetet egeegggtee geetgeteee teteegette geteetette
                                                                            540
    45 tettetttae cetteteete teteeteete tgetgetete teeteteete eegetettet
                                                                            600
    47 ctctcctcct ctcctgctct ctcctcttcc cttagctcct cttcttttct tctcctcttc
                                                                           660
    49 tteeetetee tegeetetee eetgeteete ttetetegte teeeeteeee teeegeetet
                                                                           720
    51 ctctccctc tccctctccc actcqccccg ctcqctcqct cqctqtcqca caqactcacc
                                                                           780
    53 gtcccttgtc caattatcat attcatcacc cgcaagatat caccgtgtgt gcactcgcgt
                                                                           840
    900
    57 gagagagaga gagaggeteg gteceaetge teeetgeace gegtaagtat ettettette
                                                                           960
    59 ccctcgtgag tccctcccct tttccagaat cacttgcact gtcttgttct tgaatgagaa
                                                                           1020
    61 aggaagaaaa gagcctccca ttactcagac ccgtgtaaac attattcccc ccaggagaaa
                                                                          1080
    63 atggtgttat tcaaatgaat cataataaaa tagcctctaa acagtttcta agcgggagcc
                                                                          1140
    65 teegtggaae teagegetee geteeteeea gtteetaaga gtaagtgate etettqgett
                                                                          1200
    67 ttatttcttt ctctttcctg ctggtggctg ggggtggcgg tggcgatggg ggggaggctg
                                                                          1260
    69 atgttgctgg acttgtcgct gatcttgtca ccttttgtgt actgtttctg gggtgtgagg
                                                                          1320
    71 aggegtttgc tecettteet tettteteet getetetett etcaggagag aggacegega
                                                                          1380
    73 gagggaccgg gtcgcttttt tgttcgtgga gatccccqct ttccgccaaa ccccatcctt
                                                                          1440
```

75 cegatetece caggetaaaa eteeggggee gqteeeettg teetttetet ttqtettqtt

1500

RAW SEQUENCE LISTING DATE: 02/21/2006
PATENT APPLICATION: US/10/567,630 TIME: 12:58:40

Input Set : F:\39467A.txt

Output Set: N:\CRF4\02212006\J567630.raw

```
77 tattataget geetttette eeggetette eaatttgett gteatttgea tacettteae
                                                                    1560
79 ttctcctttt ttaaccccag cagaggaccg ggaactggga ggaggagaga gggaggtggg
                                                                    1620
81 ggggcgctct gttactttcg tctcaaaacg ctgtcgaagc cgaattgtgg aaatccggct
                                                                    1680
83 tggagggag cggtgatggg tcccgggaaa cgcgcgcggc gcccctcttc cgagctcctg
                                                                    1740
85 gacccagggc tgggtcaagt tgagtagggt aaggcggcac cgggaggctc ggggggtcgc
                                                                    1800
87 gtggcggtgg gattgggaca ccagcacgag gaggaccgga ggatcgcggg ccgggtaaga
                                                                    1860
89 gtagggggtt cttgggcagc agaaatggga ggcgatgaat ctcccagcca tcgctggcag
                                                                    1920
91 actatggtgt tgggcagctt cggtctggtc tcgtctgggt ggtacctacc gttttgcccc
                                                                    1980
93 agttaggagg actggggagg gaggacagga gaggtgagag taattgttac tgggaagact
                                                                    2040
95 agtgaggagg gegggaagag ggagggaaga getgetatet tqeetqaqea qatcaggagg
                                                                    2100
97 gggacgcagt gggcgggggg agacatcacc caaagtccag tttagcaagt tgttgattct
                                                                    2160
99 tetggtgtge cagecegtta eteceeetge tgaagetgaa ggttggtgga gtgatggage
                                                                    2220
101 gtggggatgg taaaggagga gtaagtagct ttccacagac tcccaggtct ctggcccctt
                                                                     2280
2340
105 ggtggtgctt agtcttaaat agctcaagga ggcaggttgg agtgtgaaac tgctgttctt
                                                                     2400
107 ggcaacccag aaggctactc tgcctggggg aaggctggaa actcacctgc ttgttttat
                                                                     2460
109 ttttccgaga agatctgtgc tgtctccttg agcttataaa aacagaggaa gcacagggtg
                                                                     2520
111 gcctcctcgc aaagtcaagg ctagaagact cccttctcct gttctctttt ccactcatgc
                                                                     2580
2640
115 ctttcctaac ctaggtaggc agaaatctat tagcagagtg cgcatgggca qggcctgaca
                                                                     2700
117 ggtgtgttgt gtcaagaaag acaggtgcaa atttcctctg tgtctgtgtg tgtctgtaca
                                                                     2760
119 gctctagacc acaatgcttg ctcgagggtt ggagaggttt atgaatttat ggttgtcctg
                                                                    2820
121 gttaatagga ttgtctgggc taatgggaat tgggctgttg ttcttttgag ccctqccatq
                                                                    2880
123 tgagttcttg gggtggggg tgggggcaag ttggtatgtg tttgtttatt tttcttaagg
                                                                    2940
125 atattggcag tetactgetg aggetgtgte ceaggettet gtetgecagt cageceaaaq
                                                                     3000
127 cacccccact ttaggcagca ggtggaggga gactgacttt tcctttgctt cctaccagtt
                                                                    3060
129 tatgcctatc tcccaggtct gtgcttggca gagagagaga gagagagaga gaactgtcgt
                                                                     3120
3180
133 ctttggatag caatgagtgg tgtgtaactg ccaagaattc caaagtcagt ttgaaagtgt
                                                                     3240
135 tactgttgtt aaagcttatc tttttaagca tgctttctcc ttgcccagaa agaataggta
                                                                    3300
137 tgtacataaa ctctttcaag tcatatgtta aataatctca taaagtagaa tgagcctgtc
                                                                    3360
139 attgtcccag acatgtgcca aatgtcctag atatgaattt gatggagaaa gaaaatctca
                                                                     3420
141 agtacatgag aaggtaactg tgcttttcta ttctgatgca agatgtgaga agtcagttct
                                                                     3480
143 acagggaatt tettgeaaga aettetgagt attteeaaaa tgaaattttt tgtgtgtgtt
                                                                    3540
145 gagggaggaa aacgagagta ttcacattaa cttgtccatg ggttaaaaca tggacatgta
                                                                     3600
147 tatgtaatag taaaataggt gaagctaagg actgtggctt gatgtgtgag gaaagttgtt
                                                                     3660
149 gggaattcaa tgtaagcact atatctggct tcttaaaact tgacctttta aaattatctt
                                                                    3720
151 taaacagact acttctgtag actgagttgc acaggaatag gttggttggc aaatggtttt
                                                                    3780
153 tgctcattgg ctttgtgttt gggtagttat tgtttccatg aaaatgagat cgtatgtgtc
                                                                    3840
155 atttattctg tagacttcaa cattaacgtc ccccacctc ccaaacacac acacacac
                                                                    3900
157 ccaatacttt ccttggatgc ttttgaagtt ctttggtaat taaaatgtca tctatgccta
                                                                    3960
159 tgttcatttg ctttattttt aataggggtt atctgtgctt ggcacttatt gatattttat
                                                                    4020
161 gtgtccatta tgcagaattc tatttagttt aatcaccacc ttgtgggaaa aaaaagtcat
                                                                    4080
163 gcatacataa catgcatctt tgttctcact ttattcattt cctagcatca ttcctctata
                                                                    4140
165 agcagcacat gctatcttaa aacctaagct ggcttattct gtaagttgcc agacttcctc
                                                                    4200
167 tttatttgtt taaaactcaa acaggcctct tttcatgaat gtcttatatc attttaggga
                                                                    4260
169 ttgtcttgaa tttgcagtgt taatataaga agttttaggt ttcagattaa caaaagaaat
                                                                    4320
171 tataaaatgt gactgatgtt ataatatgaa aatagattgt gcatgatgta tcattatagg
                                                                    4380
173 attttaatta agtacctgtg taacttggaa aggaaccata tacataagga atttctcaga
                                                                    4440
```

RAW SEQUENCE LISTING

• 1

DATE: 02/21/2006 PATENT APPLICATION: US/10/567,630 TIME: 12:58:40

Input Set : F:\39467A.txt

Output Set: N:\CRF4\02212006\J567630.raw

175	cttattgcct	gtgcattctc	aaaggacatt	tagagagttc	aattttctgc	aaaaagaaaa	4500
		-	atttcacact	_			4560
			ttaatatttt				4620
181	ttttgcaggt	gaaaagtgtg	tggaattttt	gatgggatgg	gagaatctta	aatgaaatct	4680
183	taaatgattt	gagaagtata	ttatgacagg	aaatttaaaa	acctgataac	gcaatcttag	4740
			tgtcaagtga				4800
			atatctccaa				4860
189	gaaacaaatt	cttgaatgta	gtgaagggga	caatttaatg	attcaggggc	tacttaaatc	4920
191	agaccatctg	atttttcccc	tttgaatcac	taatttccag	attgatttga	aatattcttt	4980
193	gttaatgata	tcctatttga	aatttcataa	ccaggttgac	ccaagtagat	tagaggccca	5040
195	tacaaagatg	attttctaaa	agaagtcaag	tgtaggcttg	cacaatttct	tcaaataatt	5100
197	ttatcaacaa	agacagatca	tctaaataat	ccaagcagga	aaccatgcca	accttacact	5160
199	ctccctgcct	cataaaagat	ttgtctgaac	tatctggata	attaccgtaa	tgaaacactt	5220
201	ctttgtccag	aatctggact	ccagatagat	gcagtaaaag	ttgaatcctc	ctccccgaaa	5280
203	taacttcttt	attaaagtag	agcacttaac	cactttatac	ttcacgctgc	agtgttcctt	5340
205	tgaaattctt	tactgaaaat	tctttcctcc	taaccttaag	tcatcagttt	ccttagaatt	5400
207	ttgcatgtta	aagagaatgt	cagataattc	agatattaaa	ggagactctt	ttggagtagt	5460
209	taaaacctgt	tttgattata	cctggatgtt	tattcttcta	atatctttt	ctgggaggaa	5520
211	tctgctatgt	taagatatgc	attgtataag	aattactaaa	gcatttgtgt	aggttatata	5580
213	cgaagtgatg	caacaaaata	tttaatgatg	aaaaactcta	tatagacttt	cacattaatt	5640
215	aaagaggggt	ttacaggaat	agagtaagtg	tatccgatca	ataatacatt	tgggttcaaa	5700
217	ttctcatcag	tatttttctg	catccttgct	gatttggaca	tccaccagtg	ttgatcaaaa	5760
219	gcttcatatt	gcctagtgaa	actgaaaatt	aatgttaaaa	tgcaaatatg	atatgcatca	5820
221	ataataattg	caggtgaaac	atgatagctt	aatacatatc	ttgagaaata	aaggagttta	5880
223	aaaaatatca	atgataaagt	cattccatgg	cttcctttaa	attctgaact	ggaatatcat	5940
225	ggaagcactt	gggaaatgtt	tttaagagat	ttaatttata	ttatggtaac	gtaacagtac	6000
227	attttcttat	gtggtaaata	tattcatata	gatatcttgt	ttatgaaatg	tgatgctaat	6060
229	aaagtgctgt	gtcaaccggt	tattattatt	taatcatgcc	tatagcttcc	atgggttatg	6120
231	gttccagtgt	gtgctaccac	tatactttta	tttctaaatt	aaatctaagc	tatatggaga	6180
233	gatatattta	tttgtgccta	ttaatataat	gccttgtcct	ggattatata	atttatctta	6240
235	tttttcccat	ttgttttgtc	ttatttgtta	tgttccagct	ggacatttta	caacaagacc	6300
237	taaaagtatt	taaattcttt	tagcccaaga	cagatacaaa	tcgttattta	atctaaaaat	6360
239	gttgactgaa	atagaattac	aaaattagtt	tagtttggtg	aatatcaagg	gagttatatc	6420
241	ttgttcttaa	cagactccac	aagcatttct	ttccacctta	ggaagagcac	agccctcctc	6480
243	ttggctccag	catggggcag	ggatgcagct	gttgatacct	aggctagatg	agaggaagtg	6540
245	cagttgacgc	agaggtaaat	ggcagttgga	aaaggaagga	tgcctgggga	tgaccttgtg	6600
247	ctcatcagcg	acaccagtct	gtcctttcca	agcctctgtg	gcagagctgc	tcttcccaca	6660
249	gcaaggatgg	caggaggaaa	gtccagtttg	ggtgttaggg	tgaacaggga	gagaaaaaat	6720
251	actgcaaaaa	gtttgtttga	cattttgatt	ggagatccat	gtgctttgca	ggtgatagtc	6780
253	aagagaaaag	gatttgcata	caaatagaaa	agatgtaaaa	tttaaaaata	agggcaataa	6840
255	gctctatttt	ggggaaggtg	atatacacac	agaaaaaagt	cttccttgta	accgcccccc	6900
257	atgcaagtgt	ttctttgatt	aacagagctt	tgaaatgatt	catccttttt	cttgtctcag	6960
259	cctctccttg	ttctttctgt	catctgacag	ctaacctgat	ttatcagatc	taatgtgttt	7020
261	gtgtagtatt	tgtcactgca	tttttgtatt	cctgaaacca	attttattat	tagtgtttga	7080
			attcaatttt	-			7140
265	atctccattc	taagttcagg	aaattttatc	ctgaagcatg	caaaaagtat	ttcattctca	7200
267	agcatgcaaa	tatatata	tatatata	tatatata	tatatatata	tatatatata	7260
269	taaagaggta	tcattttgct	ttcatgatac	cctaaagcag	gctcttttaa	aatgttttat	7320
			caaagatttc				7380

RAW SEQUENCE LISTING DATE: 02/21/2006 PATENT APPLICATION: US/10/567,630 TIME: 12:58:40

Input Set : F:\39467A.txt
Output Set: N:\CRF4\02212006\J567630.raw

			gcattgagca				7440
	_	_	tggatagtca	_			7500
277	agggagcacg	tagatgtttt	tacaggtgaa	ccgaaagaga	tgggagccgt	tccagacact	7560
			tggaccctgt				7620
281	gtgaagttta	atattagata	aacccaacgt	gaaaaaaatt	ttcattttct	tcataaaatg	7680
283	ttaattataa	acaaaaagat	gtgacatctt	atatgtctac	aaaatttggg	attagcatca	7740
285	ctagttaata	agttacacaa	tgtcaagtgc	cttttatgaa	attcaaagaa	ggatgttctc	7800
287	tttttatact	gtgtttccaa	gaaacaatgg	aagttcatat	acaaagaaat	atttcccttt	7860
289	ctcacacatt	tgatggacat	tattttcttt	cttctttata	tatcttcttt	cagttttttc	7920
291	tgttttttt	tttcctttaa	tttggcacag	gaaataaggt	tcacaaatcc	tgtatgttaa	7980
293	agagtttctt	tgggcattgg	acatattatt	ttggcagatt	taaacagaag	gaaactagtc	8040
295	ctgaagatat	atttatcttt	atctcggtca	ataacttatt	attcctcata	ttgatttcta	8100
297	aaatgtggta	acatccttgt	tttgcagtga	atccaacttt	gtaataattt	gtcattaaaa	8160
299	ggacattatg	aaaatgtata	aatattctta	tagttacatt	aagatatatc	aacagatatc	8220
301	atcttcacct	atgattttac	aagtaaaaaa	tgcatagcta	agctaaataa	gcagacttat	8280
303	aaaatgacta	ttgtgcattt	atttcaatgc	taaactgacc	atttatgttt	gaaagatgct	8340
305	gctgctaagg	gtgttctcct	tcccatttta	catatgacaa	aaatattgta	aaattcaaga	8400
307	ataaaagctc	tctattatat	atttgcattt	attttagagt	ccttttcctt	taatagcgtt	8460
309	aaaaccacac	taattgtaat	gcagaaatgc	aatttttcat	gtgaatttct	catagtctca	8520
311	aaatttaacc	ttatttctta	agtatagagc	agtttcatct	tccttataat	atgaatctca	8580
313	atgcccaaaa	tttaatcaat	tggttgtcag	aggctgtgtt	cttataatct	actgtttctt	8640
315	ctgaagataa	acagtatcat	tttaggcatt	tgtgagagag	aatcatatta	ctggtgctta	8700
317	agcagttttt	gcttaatttt	tttttaatct	taatccatct	taaaccagtg	gagcagaaat	8760
319	atttaaaaat	gtttcatttc	aagcagagtg	cataataaat	tgcaataatt	gtaatgtgcc	8820
321	ataaatccca	gagcctatgc	attttgcatt	tgattcagga	ttgaggtcag	gaaatttgga	8880
323	gaaatttaaa	gaaaatgatt	catcagtcct	tttgttctgt	tggccagggt	cccgggattc	8940
			gagcttttga				9000
327	gaccatgaca	gcacagccct	cttaagccgg	caaaccaaga	ggagaagagt	tgacattgga	9060
329	gtgaaaagga	cggtagggac	agcatctgca	ttttttgcta	aggcaagagc	aacgttttt	9120
			ttctgagcag				9180
			actccgcaag				9240
			agcaaccata				9300
			cagtttccaa				9360
			cagcaactct				9420
			gagccagttt				9480
343	agagcaaagc	gcgcccgggt	tgagaatata	attcggggta	tgagccattc	ccccagtgtg	9540
345	gcattaaggg	gcaatgaaaa	tgaaagagag	atggccccgc	agtctgtgag	tccccgagaa	9600
			caagcaaaag				9660
			agaacagaag				9720
			gctgcgccag				9780
			tgaagatggt				9840
355	atcctggatg	ccagggccca	ggactctgtc	ggaaggtcag	ataatgagat	gtgcgagcta	9900
			ccgagctcga				9960
			caacaacaaa				10020
			tgagaccttg				10080
			agtcttttcg				10140
			ccaggccaga				10200
			gcagtgcttt				10260
			ggccagttcc				10320
		-		_			

RAW SEQUENCE LISTING DATE: 02/21/2006
PATENT APPLICATION: US/10/567,630 TIME: 12:58:40

Input Set : F:\39467A.txt

Output Set: N:\CRF4\02212006\J567630.raw

			ccagtctgcc			_	10380
			ctctgccacc				10440
			ggcctatcca				10500
			agcctctcct				10560
			atctcaccac				10620
			gctctccttg				10680
			accttattcg				10740
			aaggtttccc				10800
			taagaaggca				10860
			ttcctgattg				10920
			cactgccttg				10980
393	gcagtgtttt	ttttctgctg	ttcttcctgc	attatcttct	tatccctatc	tcttaaaaaa	11040
395	aataaagaag	tagatttaga	gatgagaaaa	cagtctcatt	gtaaatactg	attgaattct	11100
			ggtaagttta				11160
			gaattctcag				11220
			tcttaggaag				11280
			ggagttagtc				11340
			ctgatttatt				11400
			tggtgcagct				11460
			cccaacatac				11520
411	gcacaattaa	atataaaact	cagctgtatt	cgacttaaaa	atggctcttt	ttatgattct	11580
			tttatgtaga				11640
415	tcccgaactc	cagctatttt	taactgaaga	tttttttc	tctctgtata	tcggttcttt	11700
			ttgtggtcgt				11760
			tgtttcttat				11820
			atttaggatt				11880
			actccagagg				11940
			gttagtggag				12000
			tgaactacta				12060
429	ttaaacatgg	agtctctctt	atactgagca	agagaagaac	attgtaacag	aaagggaaga	12120
			acatatacgt				12180
			tctaaaattc				12240
435	aagcaccatg	ccttacatat	tataagcagg	cagtaaatat	tgatcaaatg	attggaatat	12300
			gcaaagtaag				12360
			tcagaggttc				12420
			ttaacttcat				12480
			ttacacacac				12540
			aattttggtg				12600
447	aactcctgta	agtccacact	aagctcttta	gttcaagatg	ccagtttatg	gtttttcttt	12660
449	aaattagact	tttcattata	accagatcat	tataattatg	gctgtgcttt	ttgtttttag	12720
451	tcttctagga	aaaaaatctt	ttagattgct	ttaagtgttg	gctatgttca	ttgtctcaac	12780
453	ctctccaaat	ccccggagga	attttgagga	tttgaattga	aataagttcc	ttttattttg	12840
			agaaaatata				12900
			tttctgtcga				12960
			cacactcaga				13020
			gattacattt				13080
			atacattgca				13140
			aggcaaacaa				13200
467	cagtgaggat	ctgttggttt	gtagagagat	tgggcccact	tttctcctgt	ctctacctct	13260

RAW SEQUENCE LISTING ERROR SUMMARY PATENT APPLICATION: US/10/567,630

DATE: 02/21/2006 TIME: 12:58:41

Input Set : F:\39467A.txt

Output Set: N:\CRF4\02212006\J567630.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:42; N Pos. 9521

VERIFICATION SUMMARY

DATE: 02/21/2006

PATENT APPLICATION: US/10/567,630

TIME: 12:58:41

Input Set : $F:\39467A.txt$

Output Set: N:\CRF4\02212006\J567630.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application No L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date L:10172 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:42 after pos.:9480